

DOC1: Hello, my name is Fred.
 DOC2: Hello, my name is Scott.
 DOC3: Scott says, "Live and let live."

1. name
 2. fred
 3. scott
 4. Live
 1 1 0 0
 1 0 1 0
 0 0 1 2

(Space required:
 16 short ints at 2 bytes each = 32 bytes)

FIG. 1
 DOCUMENT CORPUS

FIG. 2
 DICTIONARY

FIG. 3
 DENSE MATRIX - INTEGER FORMAT
 (PRIOR ART)

0.707 0.707 0.0 0.0
 0.707 0.0 0.707 0.0
 0.0 0.0 0.477 0.894

(Space required:
 12 floats at 4 bytes each
 = 48 bytes)

(Note: $0.707 = 1 * 1 / (1^2 + 1^2)^{1/2}$; $0.477 = 1 * 1 / (1^2 + 2^2)^{1/2}$; $0.894 = 2 * 1 / (1^2 + 2^2)^{1/2}$)

FIG. 4
 DENSE MATRIX - FLOATING POINT NUMBER FORMAT
 (PRIOR ART)

(1 0.707) (2 0.707)
 (1 0.707) (3 0.707)
 (3 0.477) (4 0.894)

(Space required:
 6 short ints & 6 floats
 = $6 * 2 + 6 * 4 =$ 36 bytes)

FIG. 5
 SPARSE MATRIX - FLOATING POINT NUMBER FORMAT
 (PRIOR ART)



1, 2: 0.707
1, 3: 0.707
3, 4, 4: 0.447

(Space required:
7 short ints & 3 floats
= $7*2 + 3*4 = \underline{26}$ bytes)

(Note: $0.707 = 1*1/(1^2 + 1^2)^{1/2}$; $0.447 = 1*1/(1^2 + 2^2)^{1/2}$)

FIG. 6

SMALL SPARSE MATRIX - FLOATING POINT NUMBER FORMAT

ALLDATA = 1 2 1 3 3 4 4
STARTMARKER = 1, 3, 5
MULT = 0.707 0.707 0.447

FIG. 7

SMALL SPARSE MATRIX IN VECTOR FORM

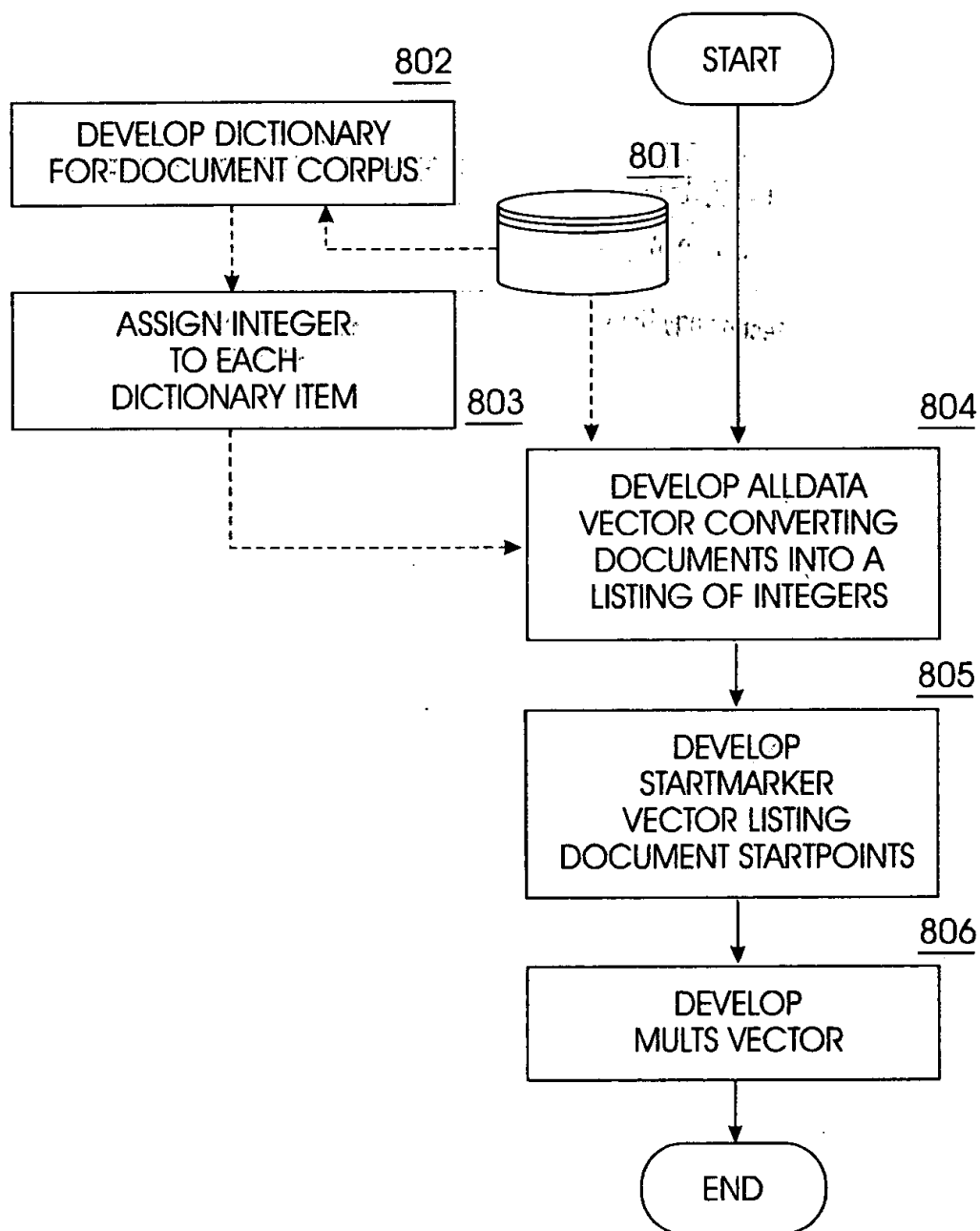


FIG. 8

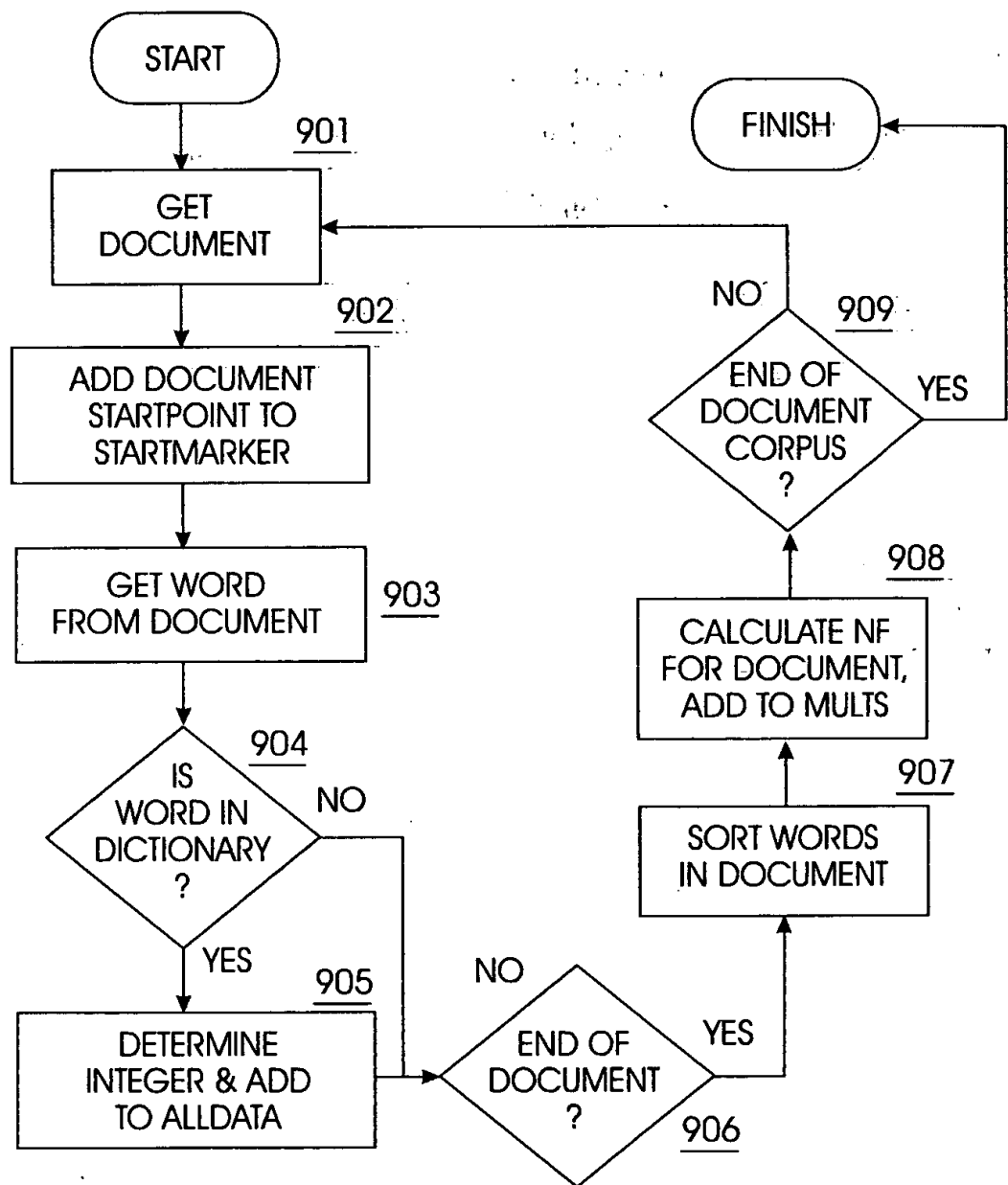


FIG. 9

